

10/649,124

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日 2 0 0 3 年 1 月 3 1 日
Date of Application:

出 願 番 号 特 願 2 0 0 3 - 0 2 5 0 7 3
Application Number:
[ST. 10/C]: [J P 2 0 0 3 - 0 2 5 0 7 3]

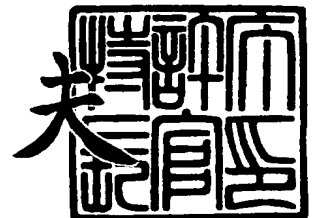
出 願 人 株式会社日立製作所
Applicant(s):



2 0 0 3 年 8 月 1 3 日

特許庁長官
Commissioner,
Japan Patent Office

今 井 康 夫



出証番号 出証特 2 0 0 3 - 3 0 6 5 1 6 0

【書類名】 特許願

【整理番号】 HI020900

【提出日】 平成15年 1月31日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 3/06

【発明者】

【住所又は居所】 神奈川県横浜市戸塚町 5 0 3 0 番地 株式会社日立製作所ソフトウェア事業部内

【氏名】 雑賀 信之

【発明者】

【住所又は居所】 神奈川県横浜市戸塚区戸塚町 5 0 3 0 番地 株式会社日立製作所ソフトウェア事業部内

【氏名】 藤本 修二

【発明者】

【住所又は居所】 神奈川県小田原市中里 3 2 2 番地 2 号 株式会社日立製作所 S A N ソリューション事業部内

【氏名】 栗倉 願讓

【特許出願人】

【識別番号】 000005108

【氏名又は名称】 株式会社日立製作所

【代理人】

【識別番号】 100071283

【弁理士】

【氏名又は名称】 一色 健輔

【選任した代理人】

【識別番号】 100084906

【弁理士】

【氏名又は名称】 原島 典孝

【選任した代理人】

【識別番号】 100098523

【弁理士】

【氏名又は名称】 黒川 恵

【選任した代理人】

【識別番号】 100112748

【弁理士】

【氏名又は名称】 吉田 浩二

【選任した代理人】

【識別番号】 100110009

【弁理士】

【氏名又は名称】 青木 康

【手数料の表示】

【予納台帳番号】 011785

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 データ管理方法、コントローラ、及びプログラム

【特許請求の範囲】

【請求項 1】 コンピュータシステムの性能情報を含むデータである性能データとそれ以外のデータとを書き込むための記憶領域と、

前記記憶領域を制御するコントローラと

を備える前記コンピュータシステムにおける前記性能データのデータ管理方法であって、

前記コントローラが、前記記憶領域の空き容量を検出するステップと、

前記コントローラが、検出した該空き容量に従って、前記性能データの書き込み方法を決定するステップと、

前記コントローラが、前記性能データを取得するステップと、

前記コントローラが、前記書き込み方法に従って、取得した前記性能データを前記記憶領域に書き込むステップと

を備えることを特徴とするデータ管理方法。

【請求項 2】 前記書き込み方法が、性能データの種類または書き込み時期であることを特徴とする請求項 1 に記載のデータ管理方法。

【請求項 3】 前記コントローラが、前記書き込み方法に従って、取得した前記性能データを前記記憶領域に書き込むステップが、

前記コントローラが、検出した前記空き容量が、設定されている容量より大きい場合、該性能データを前記記憶領域に書き込むステップと、

前記空き容量が、前記設定されている容量より小さい場合、前記記憶領域に保存されている性能データを削除し、該性能データを前記記憶領域に書きこむステップと、

を備えることを特徴とする請求項 1 または 2 に記載のデータ管理方法。

【請求項 4】 前記コントローラが、前記書き込み方法に従って、取得した前記性能データを前記記憶領域に書き込むステップが、

前記コントローラが、取得した該性能データを前記記憶領域に記録する際に必要な性能データ量を算出するステップと、

前記コントローラが、検出した前記空き容量が、設定されている容量より大きい場合、取得した該性能データを前記記憶領域に書き込み、

前記空き容量が、前記設定されている容量より小さい場合、前記性能データを記録するための性能データ領域として、前記記憶領域上に保存されている既存性能データの容量と同じ容量を割り当て、算出した該性能データ量と等しい容量の該既存性能データを削除し、取得した該性能データを前記記憶領域に書き込むステップと、

を備えることを特徴とする請求項 1 または 2 に記載のデータ管理方法。

【請求項 5】 コンピュータシステムの性能情報を含むデータである性能データとそれ以外のデータとを書き込むための記憶領域と、

前記記憶領域を制御するコントローラと

を備える前記コンピュータシステムにおける前記性能データのデータ管理方法であって、

前記コントローラが、前記記憶領域の空き容量を検出するステップと、

前記コントローラが、検出した該空き容量が第 1 の設定されている容量より大きい場合、該空き容量に従って、前記性能データの書き込み方法を決定し、前記性能データを取得し、前記書き込み方法に従って該性能データを前記記憶領域に書き込み、

前記記憶領域上の該空き容量が、第 1 の設定されている容量より小さく、前記記憶領域上に保存されている既存性能データ量が、第 2 の設定されている容量より小さい場合、前記コントローラが、前記性能データを取得し、取得した該性能データを記憶領域に記録する際に必要な性能データ量を算出し、

前記記憶領域上に保存されている該既存性能データ量と算出した該性能データ量の和が第 2 の設定されている容量より小さいと、該性能データを前記記憶領域に書きこみ、

保存されている該既存性能データ量と算出した該性能データ量との和が第 2 の設定されている容量より大きいと、少なくともこれら性能データ量の和から第 2 の設定されている容量を引いた差と同じ容量の保存されている該既存性能データを削除し、取得した該性能データを前記記憶領域に書き込み、

前記記憶領域上の該空き容量が、第 1 の設定されている容量より小さく、前記記憶領域上に保存されている該性能情報量が、第 2 の設定されている容量より大きい場合、

前記コントローラが、取得した前記性能データを記憶領域に記録する際に必要な性能データ量を算出し、

第 2 の設定されている容量中の空き容量が、少なくとも算出した該性能データ量と同じ量になるまで、保存されている該既存性能データを削除し、取得した該性能データを前記記憶領域に書き込むステップと、

を備えることを特徴とするデータ管理方法。

【請求項 6】 コンピュータシステムの性能情報を含むデータである性能データとそれ以外のデータとを書き込むための記憶領域と、

前記記憶領域を制御するコントローラと

を備える前記コンピュータシステムにおける前記性能データのデータ管理方法であって、

前記コントローラが、前記性能データを記録するために、前記記憶領域上の設定されている容量を割り当てるステップと、

前記コントローラが、前記性能データを取得するステップと、

前記コントローラが、取得した前記性能データを前記記憶領域に記録する際に必要な性能データ量を算出するステップと、

前記コントローラが、算出した該性能データ量が、前記割り当てられた容量中の空き容量より小さい場合、該性能データを前記記憶領域に書き込み、

算出した該性能データ量が、前記割り当てられた容量中の空き容量より大きい場合、前記記憶領域上に既に記録されている既存性能データを削除し、取得した該性能データを前記記憶領域に書き込むステップと、

を備えることを特徴とするデータ管理方法。

【請求項 7】 コンピュータシステムの性能情報を含むデータである性能データとそれ以外のデータとを書き込むための記憶領域と、

前記記憶領域を制御するコントローラと

を備える前記コンピュータシステムにおける前記性能データのデータ管理方法

であって、

前記コントローラが、前記性能データを記録するために、前記記憶領域上の設定されている容量を割り当てるステップと、

前記コントローラが、前記性能データを取得するステップと、

前記コントローラが、取得した該性能データを前記記憶領域に記録する際に必要な性能データ量を算出するステップと、

前記コントローラが、算出した該性能データ量が、割り当てられた前記設定されている容量中の空き容量より小さい場合、該性能データを前記記憶領域に書き込み、

算出した該性能データ量が、割り当てられた前記設定されている容量中の空き容量より大きい場合、前記記憶領域上に既に記録されている前記性能データを、その空き容量が算出した該性能データ量と同じ量になるまで削除し、取得した該性能データを前記記憶領域に書き込むステップと、

を備えることを特徴とするデータ管理方法。

【請求項 8】 前記コントローラが、前記記憶領域に保存されている前記性能情報を削除する前に、削除する前記性能データまたはその一部を含むデータのバックアップを取ることを特徴とする請求項 1, 2, 6、または 7 のいずれかに記載のデータ管理方法。

【請求項 9】 コンピュータシステムの性能情報を含むデータである性能データと、それ以外のデータとを書き込むための記憶領域を制御するためのコントローラであって、

前記記憶領域の空き容量を検出し、

検出した該空き容量に従って、次に取得する前記性能データの書き込み方法を決定し、

前記性能データを取得し、

取得した前記性能データを、前記書き込み方法に従って、前記記憶領域に書き込むことを特徴とするコントローラ。

【請求項 10】 コンピュータシステムの性能情報を含むデータである性能データと、それ以外のデータとを書き込むための記憶領域を制御するコントローラ

ラであって、

前記性能データを記録するために、前記記憶領域上の設定されている容量を割り当て、

前記性能データを取得し、

取得した前記性能データを前記記憶領域に記録する際に必要な性能データ量を算出し、

算出した該性能データ量が、割り当てられた容量中の空き容量より小さい場合、該性能データを前記記憶領域に書き込み、

算出した該性能データ量が、割り当てられた容量中の空き容量より大きい場合、前記記憶領域上に既に記録されている既存性能データを削除し、取得した該性能データを前記記憶領域に書き込む、

ことを特徴とするコントローラ。

【請求項 1 1】 コンピュータシステムの性能情報を含むデータである性能データとそれ以外のデータとを書き込むための記憶領域を制御するコントローラに、

前記記憶領域の空き容量を検出する機能と、

検出した該空き容量に従って、次に取得する前記性能データの書き込み方法を決定する機能と、

前記性能データを取得する機能と、

取得した前記性能データを、前記書き込み方法に従って、前記記憶領域に書き込む機能と

を実現させるためのプログラム。

【請求項 1 2】 コンピュータシステムの性能情報を含むデータである性能データと、それ以外のデータとを書き込むための記憶領域を制御するコントローラに、

前記性能データを記録するために、前記記憶領域上の設定されている容量を割り当てる機能と、

前記性能データを取得する機能と、

取得した前記性能データを前記記憶領域に記録する際に必要な性能データ量を

算出する機能と、

算出した該性能データ量が、割り当てられた容量中の空き容量より小さい場合、該性能データを前記記憶領域に書き込み、

算出した該性能データ量が、割り当てられた容量中の空き容量より大きい場合、前記記憶領域上に既に記録されている既存性能データを削除し、取得した該性能データを前記記憶領域に書き込む機能と、

を実現させるためのプログラム。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】

本発明は、データ管理方法、コントローラ、及びプログラムに関する。

【0 0 0 2】

【従来の技術】

コンピュータシステムは、その性能を監視することにより、性能に影響を及ぼす原因を把握し、問題発生時には迅速に対応することができるようになる。特に、近年の大規模化したコンピュータシステムに、効率よく安定した性能を発揮させるためには、その性能を監視することが重要である。例えば、処理速度の維持、データ転送速度の維持、メモリやディスクの確保、等のために、監視装置を設け、CPU、メモリ、ディスク、I/Oインターフェイス、ネットワークコントローラ等の性能を監視させ、一定時間ごとに性能情報を取得させて、性能データとして記憶装置に記録させることができる。

【0 0 0 3】

従来のコンピュータシステムの性能の監視装置は、同じ種類の性能について、同じタイミングで監視し、記録しており、他の要素に応じて監視方法などを変えするという技術は、例えば、特許文献 1 以外、ほとんど報告されていない。

【0 0 0 4】

【特許文献 1】

特開 2 0 0 1 - 3 2 5 1 2 6 号公報

【0 0 0 5】

【発明が解決しようとする課題】

コンピュータシステムの性能データは、通常ログファイルとして取得され、記憶装置に保存される。しかし、記憶装置はログファイルだけを記録しているわけではないので、ログファイルの容量が大きくなってくれば、他のデータのための記憶容量にまで影響が及ぶ。

そこで、本発明は、性能データの管理方法を制御することにより、記憶領域の効率の良い使用方法を提供することを目的とする。

【0006】**【課題を解決するための手段】**

本発明の性能データの管理方法は、コンピュータシステムの性能情報を含むデータである性能データとそれ以外のデータとを書き込むための記憶領域と、前記記憶領域を制御するコントローラとを備える前記コンピュータシステムにおける前記性能データのデータ管理方法であって、前記コントローラが、前記記憶領域の空き容量を検出するステップと、前記コントローラが、検出した該空き容量に従って、前記性能データの書き込み方法を決定するステップと、前記コントローラが、前記性能データを取得するステップと、前記コントローラが、前記書き込み方法に従って、取得した前記性能データを前記記憶領域に書き込むステップとを備えることとする。

【0007】

ストレージシステムにおいては、記憶デバイスに設定される論理ボリューム（Logical Unit）（以下、LUと記す）が記憶領域を構成し、記憶デバイス制御装置が、記憶デバイスを制御するコントローラを構成する。

【0008】**【発明の実施の形態】**

===全体構成例===

ストレージシステム100は、記憶デバイス200と、記憶デバイス制御装置300とを備えている。記憶デバイス制御装置300は、システム外の情報処理装置400から受信した入出力要求等のコマンドに従って、記憶デバイス200の制御を行う。例えば、記憶デバイス制御装置300は、情報処理装置400か

らデータ及びデータの書き込み要求を受信して、記憶デバイス200に対してデータの書き込みを行う。この際、データは、記憶デバイス200が備えるディスクドライブにより提供される物理的な記憶領域上に論理的に設定される記憶領域であるLUに記憶される。また記憶デバイス制御装置300は、情報処理装置400との間で、ストレージシステム100自体を管理するための各種コマンドの授受も行う。

【0009】

情報処理装置400はCPU (Central Processing Unit) やメモリを備えたコンピュータである。情報処理装置400が備えるCPUにより各種プログラムが実行されることにより様々な機能が実現される。情報処理装置400は、例えばパーソナルコンピュータやワークステーションであってもよく、メインフレームコンピュータであってもよい。

【0010】

図1において、情報処理装置400は、LAN (Local Area Network) 500を介して記憶デバイス制御装置300と接続しているが、SAN等他のネットワークを介してもよく、また、直接記憶デバイス制御装置300と接続していてもよい。LAN500を構成するには、インターネットを利用してもよく、専用のネットワークを構築してもよい。LAN500を介して行われる情報処理装置400と記憶デバイス制御装置300との間の通信は、例えばTCP/IPプロトコルに従って行われる。情報処理装置400からは、ストレージシステム100に対して、ファイル名指定によるデータアクセス要求（以下、ファイルアクセス要求と記す）が送信される。

【0011】

LAN500にはバックアップデバイス600が接続されている。バックアップデバイス600は、例えばMOやCD-R、DVDなどのディスク系デバイス、DATテープ、カセットテープ、オープンテープ、カートリッジテープなどのテープ系デバイスであってもよく、ハードディスクであってもよい。バックアップデバイス600は、LAN500を介して記憶デバイス制御装置300との間で通信を行うことにより、記憶デバイス200に記憶されているデータのバック

アップデータを記憶する。またバックアップデバイス 6 0 0 を情報処理装置 4 0 0 と接続してもよい。この場合は情報処理装置 4 0 0 を介して記憶デバイス 2 0 0 に記憶されているデータのバックアップデータを取得できるようにする。

【 0 0 1 2 】

=== 記憶デバイス ===

記憶デバイス 2 0 0 は多数のディスクドライブ（物理ディスク）を備えており、情報処理装置 4 0 0 に対して記憶領域を提供する。ディスクドライブとしては、例えばハードディスク装置やリムーバブルディスク装置、半導体記憶装置等様々なものを用いることができる。データは、ディスクドライブにより提供される物理的な記憶領域上に論理的に設定される記憶領域である L U に記憶されている。記憶デバイス 2 0 0 に設定される L U には、情報処理装置 4 0 0 や、L A N 上の他の端末からアクセス可能であり、それらからデータの出入力が可能である。

【 0 0 1 3 】

なお、記憶デバイス 2 0 0 は、例えば、複数のディスクドライブによりディスクアレイを構成するようにしてもよい。この場合、情報処理装置 4 0 0 に対して提供される記憶領域は、R A I D により管理された複数のディスクドライブにより提供されるようにしてもよい。

【 0 0 1 4 】

記憶デバイス制御装置 3 0 0 と記憶デバイス 2 0 0 との間は、図 1 のように、ストレージシステム 1 0 0 として直接に接続される形態としてもよく、記憶デバイス 2 0 0 をシステム外に置いて、ネットワークを介して接続するようにしてもよい。さらに記憶デバイス 2 0 0 を記憶デバイス制御装置 3 0 0 と一体として構成してもよい。

=== 記憶デバイス制御装置 ===

記憶デバイス制御装置 3 0 0 は、情報処理装置 4 0 0 との間で通信を行うための通信インタフェースを備え、情報処理装置 4 0 0 との間でデータ入出力コマンド等を授受する機能を備える。これによりストレージシステム 1 0 0 は N A S としてのサービスを情報処理装置 4 0 0 に提供することができる。

【 0 0 1 5 】

さらに、記憶デバイス制御装置 300 は、ストレージシステム 100 について、その動作状態の確認、設定や制御などを行う NAS マネージャ 700、記憶デバイス 200 の制御を行うディスク制御部 800、及び性能データを管理する性能管理プログラム 900 を備える。記憶デバイス制御装置 300 のオペレーティングシステムは例えば UNIX（登録商標）であり、これら NAS マネージャ 700 やディスク制御部 800 や性能管理プログラム 900 は、オペレーティングシステム上で、例えば、RAID マネージャ、ボリュームマネージャ、SVP マネージャ、ファイルシステムプログラム、ネットワーク制御部、バックアップ管理プログラム、障害管理プログラムのソフトウェアと共に動作する。

【0016】

NAS マネージャ 700 は、ストレージシステム 100 について、その動作状態の確認、設定や制御などを行うと共に、Web サーバとしての機能も有し、情報処理装置 400 からストレージシステム 100 の設定や制御を行うための設定 Web ページを情報処理装置 400 に提供する。NAS マネージャ 700 は、情報処理装置 400 からの HTTP（Hyper Text Transport Protocol）リクエストに応じて、設定 Web ページのデータを情報処理装置 400 に送信する。情報処理装置 400 側では、GUI（Graphical User Interface）に表示された設定 Web ページを利用して、システムアドミニストレータなどがストレージシステム 100 の設定や制御の指示を行う。NAS マネージャ 700 は、GUI における設定 Web ページに対する操作によって情報処理装置 200 から送信される設定や制御に関するデータを受信して、そのデータに対応する設定や制御を実行する。これにより、情報処理装置 400 において、ストレージシステム 100 の様々な設定や制御を行うことができる。NAS マネージャ 700 の設定 Web ページを利用して行うことができる内容としては、例えば、記憶デバイスの管理や設定（容量管理や容量拡張・縮小、ユーザ割り当て等）、上述の複製管理やリモートコピー（レプリケーション）等の機能に関する設定や制御（複製元の LU と複製先の LU の設定など）、OS や OS 上で動作するアプリケーションプログラムのバージョン管理、ウイルス検知プログラムやウイルス駆除などのデータの安全性に関する機能を提供するセキュリティ管理プログラムの動作状態の管理や設定など

がある。

【0017】

ディスク制御部 800 は、記憶デバイス 200 の制御を行い、例えば、記憶デバイス制御装置 300 が情報処理装置 400 からデータ及びデータ書き込みコマンドを受信すると、ディスク制御部 800 は、そのデータ書き込みコマンドに従って、記憶デバイス 200 へデータの書き込みを行う。また、論理アドレス指定による LU へのデータアクセス要求を、物理アドレス指定による物理ディスクへのデータアクセス要求に変換する。さらに、記憶デバイス 200 における物理ディスクが RAID により管理されている場合には、RAID 構成（例えば、RAID 0, 1, 5）に従って、データへアクセスする。またディスク制御部 800 は、記憶デバイス 200 に記憶されたデータの複製やバックアップの際の制御を行う。さらにディスク制御部 800 は、災害発生時のデータ消失防止（ディザスタリカバリ）などを目的としてプライマリサイトのストレージシステム 100 のデータの複製をセカンダリサイトに設置された他のストレージシステム 100 にも記憶させる制御（レプリケーション機能、又はリモートコピー機能）なども行う。

【0018】

性能管理プログラム 900 は、例えば、図 2 の取得項目に示すような、コンピュータシステムの性能データを監視し、後述するポリシー定義（図 6）に従って、性能データを取得する。そして、例えばログファイルの形式で、ディスク制御部を通じて、取得した性能データを記憶デバイス 200 に書き込む。

【0019】

===管理端末===

管理端末 1000 はストレージシステム 100 を保守・管理するためのコンピュータである。管理端末 1000 を操作することにより、例えば記憶デバイス 200 内の物理ディスク構成の設定や、LU の設定等を行うことができる。ここで、記憶デバイス 200 内の物理ディスク構成の設定としては、例えば物理ディスクの増設や減設等を行うことができる。さらに管理端末 1000 からは、ストレージシステム 100 の動作状態の確認や故障部位の特定、オペレーティングシス

テムのインストール等の作業を行うこともできる。また管理端末1000はLANや電話回線等で外部保守センタと接続されており、管理端末1000を利用してストレージシステム100の障害監視を行ったり、障害が発生した場合に迅速に対応することも可能である。障害の発生は例えばOSやアプリケーションプログラム、ドライバソフトウェアなどから通知される。この通知はHTTPプロトコルやSNMP (Simple Network Management Protocol)、電子メールなどにより行われる。これらの設定や制御は、管理端末1000で動作するWebサーバが提供するWebページをユーザインタフェースとしてオペレータなどにより行われる。オペレータ等は、管理端末1000を操作して障害監視する対象や内容の設定、障害通知先の設定などを行うこともできる。

【0020】

管理端末1000は記憶デバイス制御装置300に内蔵されている形態とすることもできるし、外付けされている形態とすることもできる。また管理端末1000は、記憶デバイス制御装置300及び記憶デバイス200の保守・管理を専用に行うコンピュータとすることもできるし、汎用のコンピュータに保守・管理機能を持たせたものとすることもできる。

【0021】

===実施の形態 1===

本実施の形態では、取得した性能データを記憶する記憶デバイス200の空き容量を検出し、その空き容量に従って、取得する性能データの書き込み方法を決定した後、取得した性能データを記憶デバイス200に書き込むようにする（以下、この形態を「性能データ領域変動タイプ」と呼ぶ）。このことにより、記憶デバイス200の空き容量を考慮に入れながら、性能データの記憶デバイス200における占有容量を決定することができるようになる。性能データの重要性と他のデータの重要性との大小は様々であるが、例えば、記憶デバイス200の空き容量が多く、性能データ量が多少増えても、他のデータのための容量は十分あると考えられるような場合、性能データを頻繁に、あるいは多量に取得し、保存する方が、コンピュータシステムの性能の管理には効果的である。しかしながら、記憶デバイス200の空き容量が少なく、性能データの記憶デバイス200へ

の保存が他のデータののための記憶容量に影響を及ぼしかねない場合、保存する性能データ量をできるだけ減らす必要がある。このように、本実施の形態では、性能データだけでなくユーザデータなど他のデータも記憶しなければならない記憶デバイス 200 の効率的な利用を可能にする。実施の形態 1 では、記憶デバイス 200 の空き容量に従って、性能データの取得項目数、取得項目内容及び書き込み間隔を決定する。

【0022】

まず、情報処理装置 400 において、図 2 の取得レベル定義、図 3 の取得項目、図 5 の初期ポリシー定義、図 6 のポリシー定義を規定する。取得レベル定義（図 2）では、CPU、メモリ、ディスク性能、ネットワーク性能など、コンピュータシステムの性能の中から、大きく分けて、どのような種類のものを対象にするか、についての組み合わせを定義し、各組み合わせに対し、取得レベルを規定する。取得項目定義（図 3）は、取得レベル定義において決定した項目を、さらに小項目に分けて定義を行い、情報処理装置 400 において、具体的にどの性能を対象にするかの設定を選択することができるようにする。このようにしてできた出力ファイルの例を図 4 に示す。

【0023】

初期ポリシー定義（図 5）では、性能データの初期書き込み時に、処理時間内に取得する全性能データの予測必要容量が記憶デバイス 200 の空き容量を超えた場合に用いられ、そこでは、記憶デバイス 200 に既に記憶されている性能データの処理方法が定義されている。具体的には、既に記憶されている性能データに対する処理方法として、「上書き」と「削除」から選択できる。「上書き」は、最も古いファイルから、新しく書き込む容量だけ削除し、そこに新しいファイルを書き込む方法で、「削除」は、古いファイルを全て削除する方法である。ポリシー定義（図 6）は、定常状態で、その状態の定義と、その状態になった場合にとるアクション、即ち性能データをどのように書き込むかという方法を定義している。記憶デバイス 200 の容量が、あらかじめ設定された容量より少なくなった時、記憶デバイス 200 の運用は、「停止運用ポリシー」で制御される。その制御方法を図 7 に示す。まず、性能データ領域の容量を固定させるか、変動さ



せるかの選択ができる。本実施の形態は、変動させる方を選択する場合（図 7 における 1 ～ 3 番）に相当する。さらに性能データの書き込み方に関して、「ラップアラウンド（1， 2 番）」と「削除（3 番）」とから選択できる。「ラップアラウンド」というのは、性能データ領域に特定の容量を割り当て、初めは順次書き込みをし、空き領域がなくなったら、古いファイルから順次消去しながら、新しいファイルを書き込んでいく、という方法をいう。過去の性能データファイルがある程度残しておくため、必要な時には回収できるという利点がある。また、「削除」というのは、その時点で書き込まれている性能データをすべて削除した後、新しいファイルを書き込む方法である。記憶領域の空き領域があまりにも少ない場合、過去の性能データファイルを残しておく、他のデータに対する記憶領域が残らなくなるかもしれない、そのような場合は、性能データファイルを保存せず、全部削除することにより、過去の性能データファイルから記憶容量を解放するのがよい。

【 0 0 2 4 】

以下、開始時刻から終了時刻までに性能管理プログラム 9 0 0 の行う処理手順について、図 8 のフローチャートを用いて説明する。性能管理プログラム 9 0 0 は、図 2 の取得レベル定義、図 3 の取得項目、図 5 の初期ポリシー定義、図 6 のポリシー定義を読み込み (S102)、処理開始時刻になるまで、待機する (S104)。処理開始時刻になると、性能データを書き込む記憶デバイス 2 0 0 の空き容量を検出する (S106)。この空き容量に、処理時間内に取得する全性能データの書き込みができるかどうか予測するために、例えば、次式の計算を行うことにより、処理時間内に取得する全性能データの予測必要容量を算出する (S108)。

【 0 0 2 5 】

(取得回数) X (各項目の平均容量) X (取得項目数)

各パラメーターは、アルゴリズムによって、過去に行った処理のデータから推定されてもよく、またオペレータによって外部入力される変数としてもよく、管理端末から入力されてもよい。なお、全く初めて処理する場合は、記憶デバイス 2 0 0 に何も書かれていないので、(S106) から (S112) の処理を省略する。記憶デバイス 2 0 0 の空き容量と予測必要容量を比較し (S110)、予測必要容量が記憶

デバイス 200 の空き容量より大きい場合、そのまま処理を続けると、性能データが書き込めない事態に陥ることが予想されるため、あらかじめ性能データを減らすため、図 5 の初期ポリシー定義において選択した方法に従い、既にかき込まれている性能データを処理する (S112)。一方、予測必要容量が記憶デバイス 200 の空き容量より小さい場合、既にかき込まれている性能データはそのままにしておく。いずれの場合でも、その後、必要最小限の性能データを取得し (S114)、記憶デバイス 200 にかき込む (S116)。次に、記憶デバイスの空き容量を検出し (S118)、ポリシー定義に基づき、その空き容量に対応した運用ポリシーを決定する (S120)。その運用ポリシーに従って、性能データを取得し (S122)、記憶デバイス 200 にかき込む (S124)。性能データを取得する前に終了時刻が来ると (S126)、処理時間中に取得した性能データから、性能サマリデータを作製し (S128)、処理を終了する。

【0026】

運用ポリシーにおいては、記憶デバイス 200 の空き容量が多いほど、多くの種類の性能データを取得し、書き込み間隔も短くする。しかし、空き容量が少ないと、ユーザデータなど他のデータへの侵食を防ぐため、取得する性能データの種類を減らし、書き込み間隔も長くする。

【0027】

上記処理中、初期ポリシー定義に基づく処理の詳細を、図 9 のフローチャートに示す。初期ポリシー定義で「上書き」が選択されていれば、最も古いファイルから、新しく書き込む容量分を削除して、そこに新しいファイルを書き込むことができるように、書き込み位置を性能データの最前部に置く (S202)。初期ポリシー定義で「削除」が選択されていれば、既存の性能データを全て削除する (S204)。

【0028】

さらに、ポリシー定義に基づく処理を、図 10 のフローチャートに示す。基本的には、記憶デバイス 200 の空き容量に従って、取得レベル及び性能データの書き込み間隔を決定する (S302-S308) が、空き容量が、あらかじめ設定された、ある一定値以下 (又は未満) である場合、例えば図 10 においては、1 M 未満で



ある場合に、過去に記憶された既存の性能データの処理が行われ（以下、停止運用ポリシーと呼ぶ。）、取得する性能データの取得レベルは最低レベルに抑えられ、書き込み間隔は最長になるように設定される。

【 0 0 2 9 】

停止運用ポリシーにおいて、あらかじめ「ラップアラウンド」が選択されている場合、性能データ領域に、あらかじめ設定される容量を割り当て、初めは順次書き込みをし、性能データ領域に空き領域がなくなったら、性能データの古いファイルから順次消去しながら、新しいファイルを書き込んでいくのであるが、この場合、性能データ領域の容量を、その時点で占有している容量に規定する（S310；変動タイプ）か、予め決められたサイズの容量に規定するか（S314；固定タイプ）、の2通りがある。変動タイプの場合、新しい性能データを書き込む時に、そのファイルと同じ容量だけ古いファイルから削除し（S312）、新しい性能データを書き込む（S124）。この方式には、過去の性能データファイルを、比較的長く保存できるという利点がある。一方、固定タイプの場合、あらかじめ性能データ領域の容量を設定し（S314）、新しい性能データを書き込む時に、設定した容量まで既存の性能データを削除し、さらに最低レベルの性能データファイルと同じ最低容量分だけ、古いファイルから削除し（S316）、新しい性能データファイルを書き込む（S124）。この方法には、記憶デバイス 2 0 0 の空き容量が減った時でも、性能データを削除することにより、記憶デバイス上で、性能データファイルの占める領域を、ある程度抑制することができるという利点がある。停止運用ポリシーにある間は、これら性能データ領域として割り当てられた特定の容量は維持されるが、ユーザデータなどの他のデータが削除され、記憶デバイスの空き容量が増え、停止運用ポリシーから脱した場合、ここで規定された性能データ領域の容量の固定は解除される（S320）。一方、あらかじめ「上書き」が設定されている場合、既存の性能データを全て消去する（S318）。

【 0 0 3 0 】

本実施例では、空き容量検出と運用ポリシーの決定、データの取得、データの書き込み、というステップを、この順で一連に起こるステップとして述べたが、本来これらは独立のステップであり、この順序が変わっても、各ステップ間に多



少の時差が生じて問題ない。なお、データの取得方法は、本実施例では、性能データの取得項目の種類と書き込み間隔を例にしたが、それらの組み合わせでも良く、また、それらには限らない。

【0031】

性能管理プログラムに、既存のデータを削除する際あるいはそれ以前に、削除するデータのバックアップを取るためのアルゴリズムを設けてもよい。削除する前段階でバックアップを取っても良いし、一部の処置段階でのみバックアップを取ってもよい。さらに、バックアップ全部または一部について、性能サマリデータのバックアップを取るアルゴリズムとしてもよい。バックアップは、バックアップディスクに行われてもよく、他のNASに送信されてもよい。

【0032】

===実施の形態 2===

本実施の形態では、記憶デバイス200において、あらかじめ、性能データのための容量を固定する（以下、この形態を「性能データ領域固定タイプ」と呼ぶ）。このことにより、性能データが増え続けて、記憶デバイスのうちでユーザの使用するユーザ領域が圧迫されるような状況は回避することができるようになる。

【0033】

ストレージシステム100の構成は実施例1と同じなので、ここでは説明を省略する。情報処理装置400において、ポリシー定義（図6）については、まず、図7の表に示すように、性能データ領域の容量を固定させるか、変動させるかの選択ができる。本実施の形態は、固定させる方を選択する場合に相当する（図7、表の4，5番）。

【0034】

以下、本実施の形態における、性能管理プログラム900による性能データの管理の流れを図11のフローチャートを用いて説明する。まず、性能管理プログラム900が、性能データ領域の容量を確保し、固定する(S402)。性能データを取得し(S404)、その性能データを記憶デバイスに記憶させるのに必要なデータ容量を算出する(S406)。ここで、性能データ領域の書き込み方法として、「ラッ



プアラウンド」と「削除」が選択可能である。「ラップアラウンド」が選択されている場合、性能データ領域に、初めは順次書き込みをし(S408)、性能データ領域に空き領域がなくなったら、性能データの古いファイルから順次消去しながら(S410)、新しいファイルが書き込まれる(S408)。「削除」が選択されている場合、初めは順次書き込みをし(S408)、性能データ領域に空き領域がなくなったら、既存の性能データを全て消去し(S412)、新たに性能データを書き込んでいく(S408)。処理終了時刻が来ると、記憶されている性能データから、性能サマリデータを作成し(S414)、処理を終了する。

性能サマリデータの取り扱いに関しては、実施の形態1と同じであるので、ここでは説明を省略する。

【0035】

===性能サマリデータ===

性能サマリデータとは、性能データのうちの一部を抽出したデータあるいは性能データから平均値などを算出したデータであり、性能データ全体の特徴部分を捉えたものである。従って、性能データのうち、設定されている回数ごとにサマリデータとして保存しても良く、処理速度などの数値で表れるような性能であれば、設定されている回数ごとに平均値を取り、それを保存しても良い。また、CPUの性能データのみを保存するというように、一部のデータを抽出しても良い。さらに、コンピュータシステムにかかる負荷の大きさに伴って、サマリデータの内容を変えたり、保存する頻度を増やしたりしても良い。

【0036】

例えば、実施の形態1及び2では、外部LAN500を通じて他の特定のNASに送信しておくことにより、受信する側のNASでは、全てのNASの性能サマリデータが揃うことになり、情報処理装置400から、そのNASの備える記憶デバイス200にアクセスするだけで、全ての記憶デバイスの性能サマリデータを入手でき、全ての記憶デバイスの性能を容易に比較できるようになる。従って、この性能サマリデータを監視しておけば、ある記憶デバイス200に問題が生じた時であっても、容易に発見できるようになる。

【0037】

他の実施の形態として、チャンネル制御部 350 を備える記憶デバイス制御装置 300 がある (図 12)。記憶デバイス制御装置 300 は、チャンネル制御部 350 により LAN 450 を介して情報処理装置 400 との間で通信を行う。チャンネル制御部 350 は、情報処理装置 400 からのファイルアクセス要求を個々に受け付ける。すなわち、チャンネル制御部 350 には、個々に LAN 450 上のネットワークアドレス (例えば、IP アドレス) が割り当てられていてそれぞれが個別に NAS として振る舞い、個々の NAS があたかも独立した NAS が存在しているかのように NAS のサービスを情報処理装置 400 に提供することができる。このように 1 台のストレージシステム 100 に個別に NAS としてのサービスを提供するチャンネル制御部 350 を備えるように構成したことで、従来、独立したコンピュータで個別に運用されていた NAS サーバが一台のストレージシステム 100 に集約されて運用される。そして、この構成によってストレージシステム 100 の統括的な管理が可能となり、各種設定・制御や障害管理、バージョン管理といった保守業務の効率化が図られる。

【0038】

このような構成のストレージシステム 100 において、複数のチャンネル制御部に共有される記憶デバイス 200 を設けることができる。この場合、チャンネル制御部 350 ごとの性能データを取得し、各チャンネル制御部 350 固有の記憶デバイス 200 に記憶されるが、性能サマリデータは、共有される記憶デバイスに送信され、記憶されてもよい。また、例えば、一つのチャンネル制御部 350 に複数の性能サマリデータを集約させるような場合、内部 LAN 450 に通じて、性能サマリデータを送受信することができる。なお、このような構成に於いても、外部 LAN 500 を通じて、性能サマリデータを送受信してもよい。

【0039】

【発明の効果】

本発明によって、性能データの管理方法を制御することにより、記憶領域の効率の良い使用方法を提供することができるようになる。

【図面の簡単な説明】

【図 1】

本発明の一実施の形態におけるストレージシステムの全体構成を示す図である。

。

【図 2】

本発明の一実施の形態における取得レベル定義を示す表である。

【図 3】

本発明の一実施の形態における取得項目の詳細を示す表である。

【図 4】

本発明の一実施の形態における性能データを含む出力ファイルの例である。

【図 5】

本発明の一実施の形態における初期ポリシー定義を示す表である。

【図 6】

本発明の一実施の形態におけるポリシー定義を示す表である。

【図 7】

本発明の一実施の形態における停止運用ポリシーにおける性能データ領域の選択可能な制御方法を示す表である。

【図 8】

本発明の一実施の形態において性能管理プログラムを実行した際のストレージシステムの動作を表すフローチャートである。

【図 9】

本発明の一実施の形態における初期ポリシー決定時のストレージシステムの動作を表すフローチャートである。

【図 10】

本発明の一実施の形態における運用ポリシー決定時のストレージシステムの動作を表すフローチャートである。

【図 11】

本発明の別の実施の形態において性能管理プログラムを実行した際のストレージシステムの動作を表すフローチャートである。

【図 12】

本発明の別の実施の形態において、チャンネル制御部を備える記憶デバイス制御

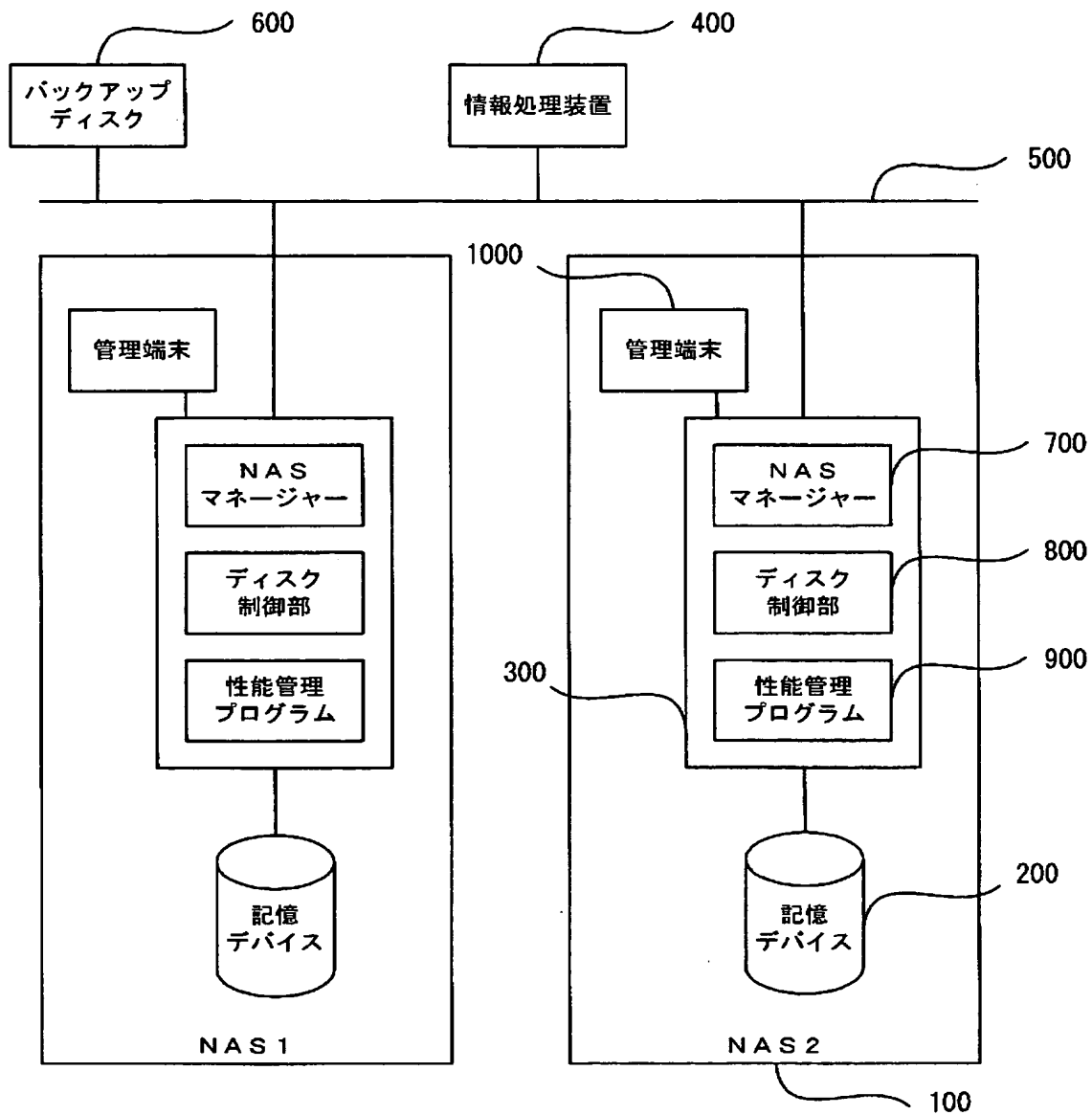
装置を用いる場合のストレージシステムの全体構成を示す図である。

【符号の説明】

1 0 0	ストレージシステム
2 0 0	記憶デバイス
3 0 0	記憶デバイス制御装置
3 5 0	チャネル制御部
4 0 0	情報処理装置
4 5 0	内部 L A N
5 0 0	外部 L A N
6 0 0	バックアップディスク
7 0 0	N A S マネージャ
8 0 0	ディスク制御部
9 0 0	性能管理プログラム
1 0 0 0	管理端末

【書類名】 図面

【図 1】



【図 2】

取得レベル	取得項目列
1	CPU、メモリ、ディスク性能、ネットワーク性能
2	CPU、メモリ、ディスク性能、
....	
....	
N	CPU

【図 3】

大分類	小分類	選択
CPU	アイドル状態の割合	x
	システムレベルの実行に使用した割合	x
	ユーザレベルの実行に使用した割合	x
	I/O待ちによるアイドル状態の割合	
メモリ	ページフォールト数/s	x
ディスク性能 (ファイルシステム)	ディレクトリ検索数/s	x
	論理入出力数/s	x
	ブロック入出力数/s	x
	Iノード検索ルーチンコール数/s	
ネットワーク性能	IPCメッセージ・プリミティブ数/s	x
	IPCセマフォ・プリミティブ数/s	x

【図 4】

ホスト名: Host001 開始時刻: YYYY/MM/DD HH:MM:SS 終了時刻: yyyy/mm/dd hh/mm/ss		
大項目	中項目	性能データ
CPU	アイドル状態	40回/s
	I/O 待ち	15回/s

メモリ	ページフォールト	3回/s

ディスク性能	ディレクトリ検索	25回/s
	論理 I/O	10回/s

ネットワーク性能	IPCメッセージ	20回/s

...

【図 5】

	初期ポリシー
1	上書き
2	削除

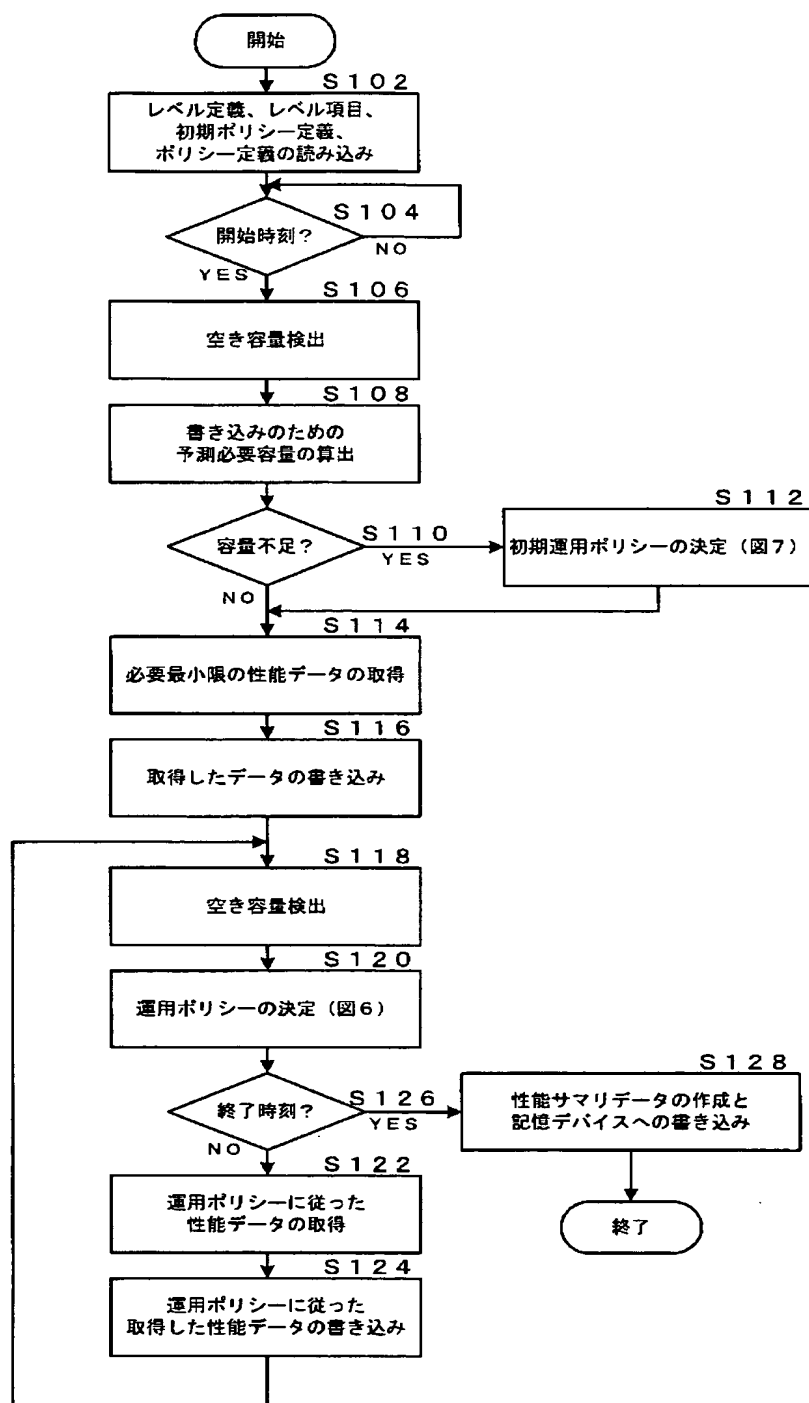
【図 6】

	残容量	取得レベル	取得間隔(秒)
停止運用ポリシー	<1M	N	1200
運用ポリシー	<100M	N	1200
	<200M	N-1	800
	<400M	N-2	700
	≥400M	K	600

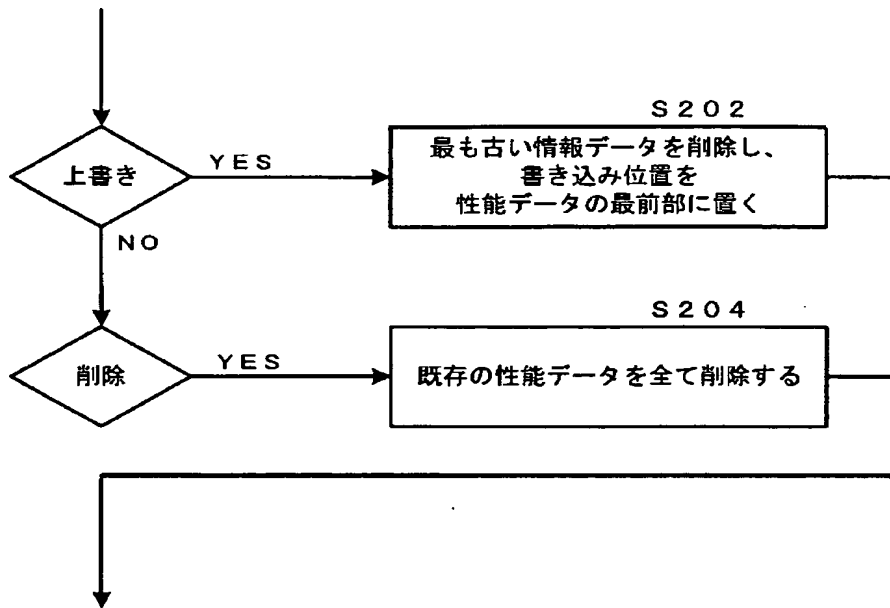
【図 7】

番号	性能データ領域のタイプ	書き込み方法	停止運用時のファイルサイズ
1	変動	ラップアラウンド	固定
2	変動	ラップアラウンド	変動
3	変動	削除	
4	固定	ラップアラウンド	
5	固定	変動	

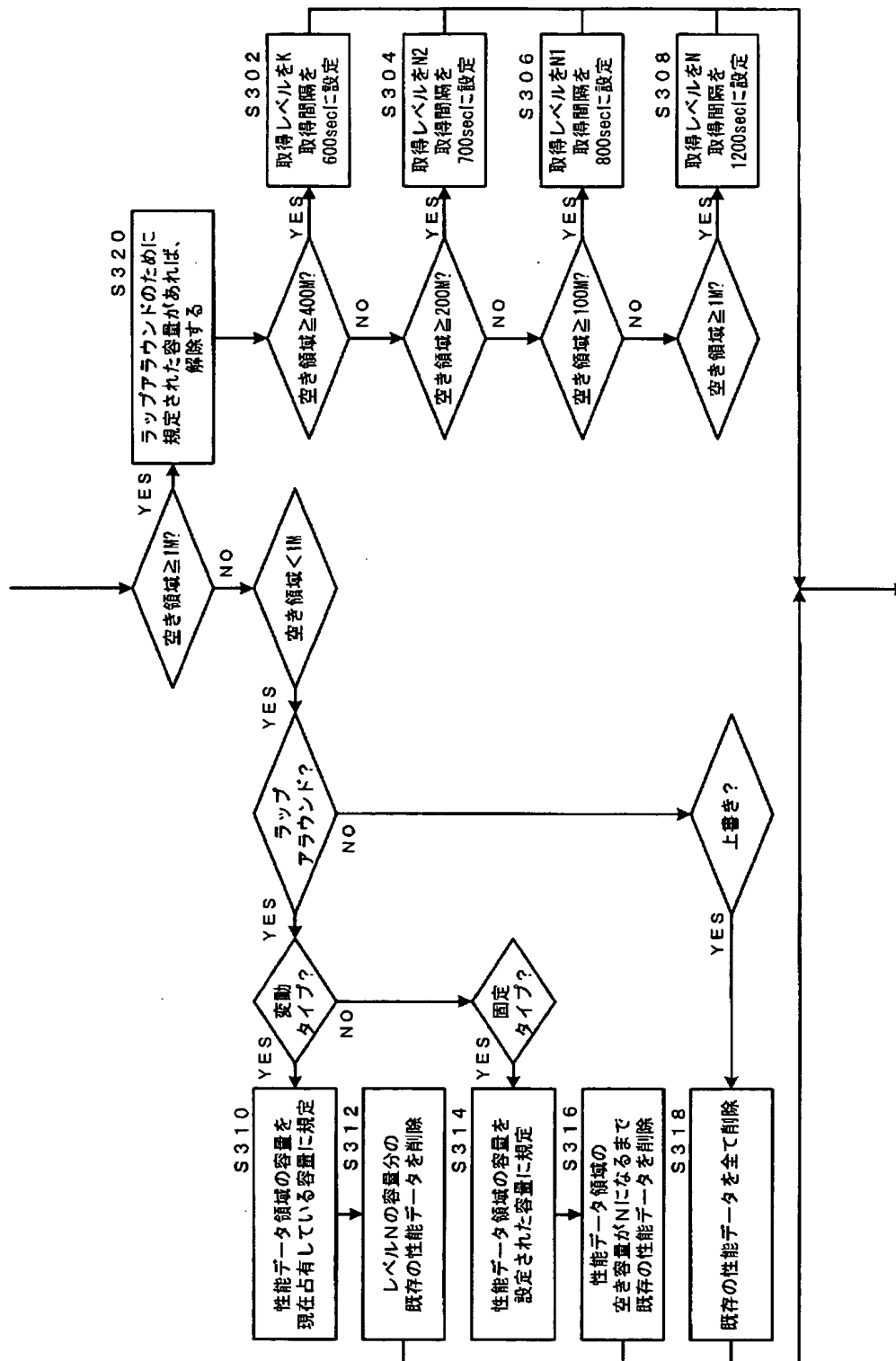
【図 8】



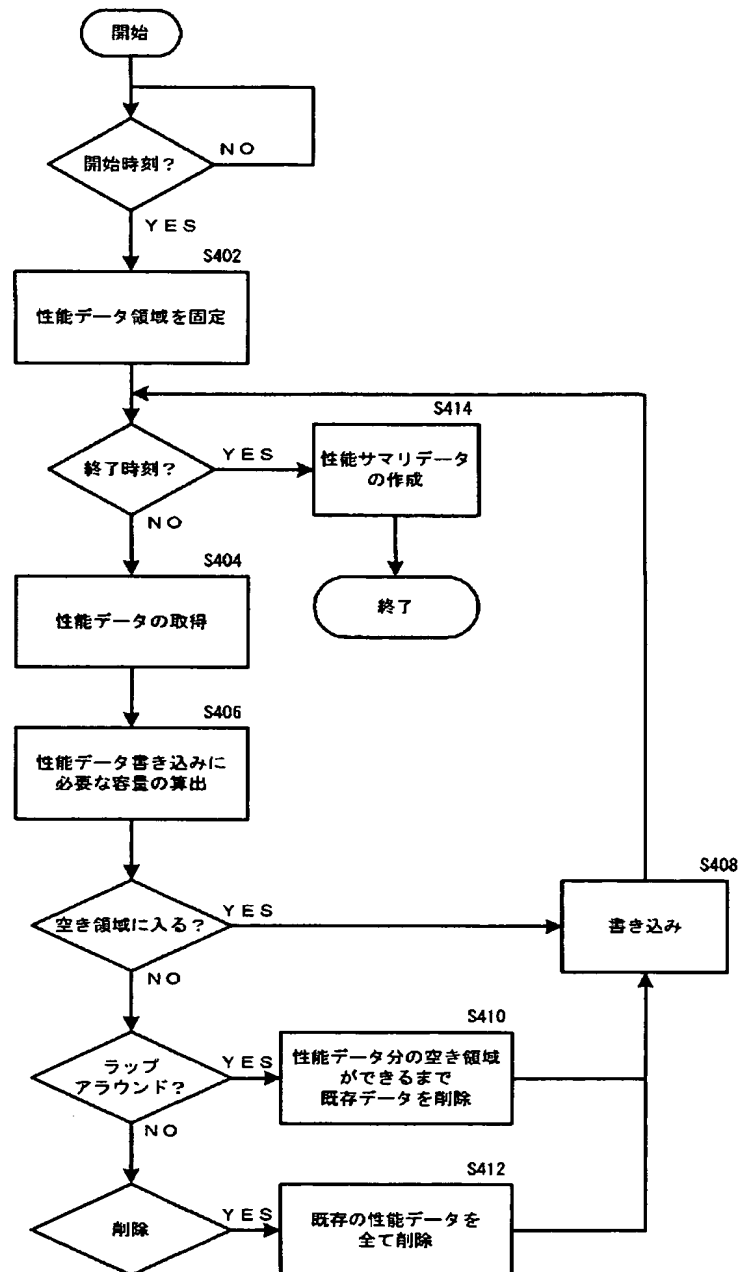
【図 9】



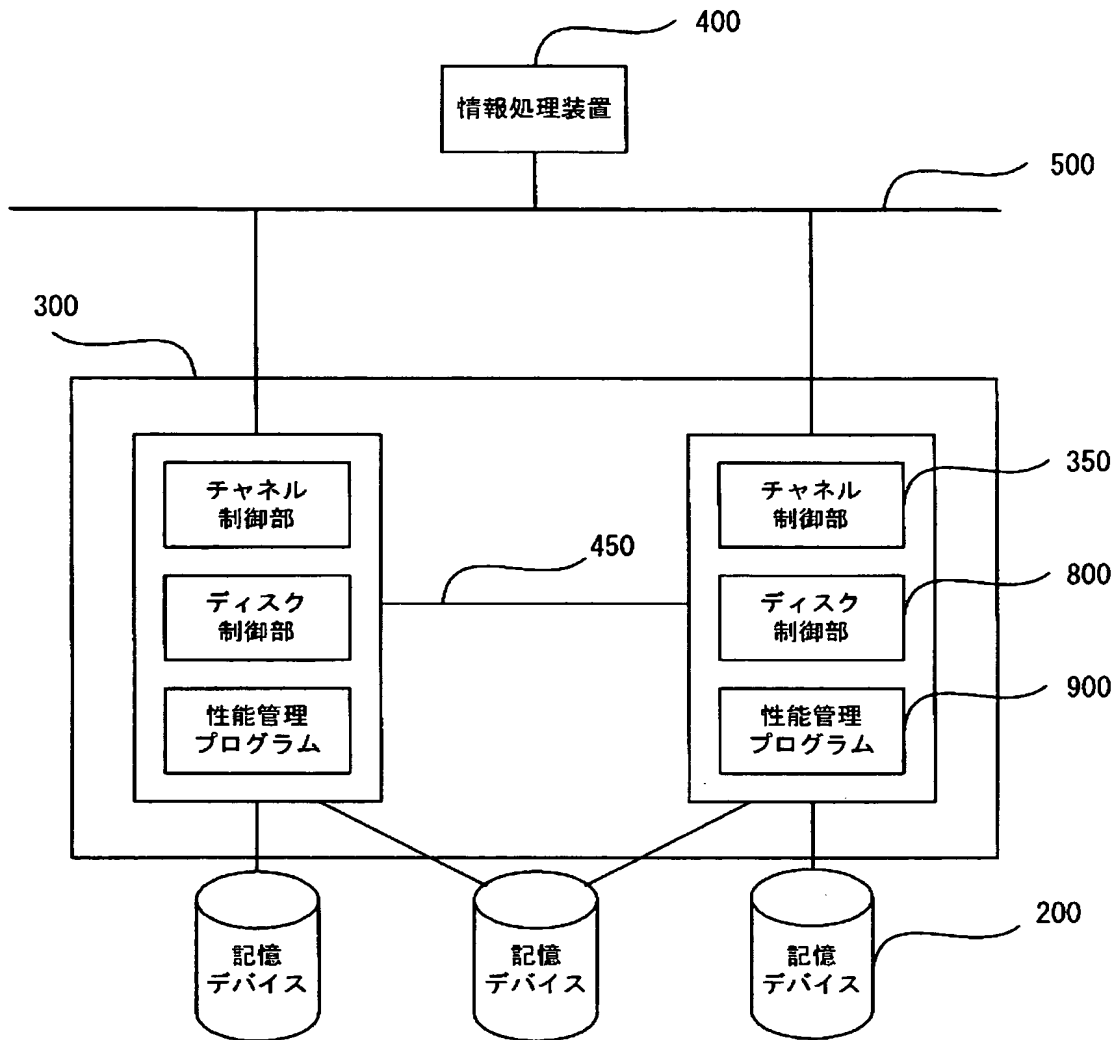
【図 10】



【図 11】



【図 12】



【書類名】 要約書

【要約】

【解決手段】 コンピュータシステムの性能情報を含むデータである性能データとそれ以外のデータとを書き込むための記憶領域と、記憶領域を制御するコントローラと、を備える前記コンピュータシステムにおける前記性能データのデータ管理方法であって、コントローラが、記憶領域の空き容量を検出するステップと、コントローラが、検出した空き容量に従って、性能データの書き込み方法を決定するステップと、コントローラが、性能データを取得するステップと、コントローラが、書き込み方法に従って、取得した性能データを記憶領域に書き込むステップとを備えるようにする。

【選択図】 図 8

特願 2003-025073

出願人履歴情報

識別番号

[000005108]

1. 変更年月日

1990年 8月31日

[変更理由]

新規登録

住 所

東京都千代田区神田駿河台4丁目6番地

氏 名

株式会社日立製作所